

Antisymmetry, pseudospectral methods, and conservative PDEs

Robert I. McLachlan and Nicolas Robidoux

Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand
 R.McLachlan@massey.ac.nz, N.Robidoux@massey.ac.nz

Abstract. “Dual composition”, a new method of constructing energy-preserving discretizations of conservative PDEs, is introduced. It extends the summation-by-parts approach to arbitrary differential operators and conserved quantities. Links to pseudospectral, Galerkin, antialiasing, and Hamiltonian methods are discussed.

1. Introduction. For all $u, v \in C^1([-1, 1])$,

$$\int_{-1}^1 v \partial_x w \, dx = - \int_{-1}^1 w \partial_x v \, dx + [vw]_{-1}^1,$$

so the operator ∂_x is skew-adjoint on $\{v \in C^1([-1, 1]) : v(\pm 1) = 0\}$ with respect to the L^2 inner product $\langle \cdot, \cdot \rangle$. Take n points x_i , a real function $v(x)$, and estimate $v'(x_i)$ from the values $v_i := v(x_i)$. In vector notation, $\mathbf{v}' = D\mathbf{v}$, where D is a differentiation matrix. Suppose that the differentiation matrix has the form $D = S^{-1}A$, in which S induces a discrete approximation

$$\langle \mathbf{v}, \mathbf{w} \rangle_S := \mathbf{v}^T S \mathbf{w} \approx \int vw \, dx = \langle v, w \rangle,$$

of the inner product. Then

$$(1) \quad \langle \mathbf{v}, D\mathbf{w} \rangle_S + \langle D\mathbf{v}, \mathbf{w} \rangle_S = \mathbf{v}^T S S^{-1} A \mathbf{w} + \mathbf{v}^T A^T S^{-T} S \mathbf{w} = \mathbf{v}^T (A + A^T) \mathbf{w},$$

which is zero if A is antisymmetric (so that D is skew-adjoint with respect to $\langle \cdot, \cdot \rangle_S$), or equals $[vw]_{-1}^1$ if $x_1 = -1$, $x_n = 1$, and $A + A^T$ is zero except for $A_{nn} = -A_{11} = \frac{1}{2}$. Eq. (1) is known as a “summation by parts” formula; it affects the energy flux of methods built from D . More generally, preserving structural features such as skew-adjointness leads to natural and robust methods.

Although factorizations $D = S^{-1}A$ are ubiquitous in finite element methods, they have been less studied elsewhere. They were introduced for finite difference methods in [4] (see [9] for more recent developments) and for spectral methods in [2], in which the connection between spectral collocation and Galerkin methods was used to explain the skew-adjoint structure of some differentiation matrices.

Let $\mathcal{H}(u)$ be a continuum conserved quantity, the *energy*. We consider PDEs

$$(2) \quad \dot{u} = \mathcal{D}(u) \frac{\delta \mathcal{H}}{\delta u},$$

and corresponding “linear-gradient” spatial discretizations [5, 6, 7], ODEs of the form

$$(3) \quad \dot{\mathbf{u}} = L(\mathbf{u}) \nabla H(\mathbf{u})$$

with appropriate discretizations of u , \mathcal{D} , \mathcal{H} , and $\delta/\delta u$. For a PDE of the form (2), if $\mathcal{D}(u)$ is formally skew-adjoint, then $d\mathcal{H}/dt$ depends only on the total energy flux through the boundary; if this flux is zero, \mathcal{H} is an integral. Analogously, if (3) holds, then $\dot{H} = \frac{1}{2}(\nabla H)^T (L + L^T) \nabla H$, so that H cannot increase if the symmetric part of L is negative

definite, and H is an integral if L is antisymmetric. Conversely, all systems with an integral can be written in “skew-gradient” form ((3) with L antisymmetric) [7]. Hamiltonian systems are naturally in the form (2) and provide examples.

This paper summarizes [8], which contains proofs and further examples.

2. Discretizing conservative PDEs. In (2), we want to allow constant operators such as $\mathcal{D} = \partial_x^n$ and $\mathcal{D} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and nonconstant ones such as $\mathcal{D}(u) = u\partial_x + \partial_x u$. These differ in the class of functions and boundary conditions which make them skew-adjoint, which suggests Defn. 1 below.

Let $(\mathfrak{F}, \langle, \rangle)$ be an inner product space. We use two subspaces \mathfrak{F}_0 and \mathfrak{F}_1 which can be infinite dimensional (in defining a PDE) or finite dimensional (in defining a discretization). We write $\{f_j\}$ for a basis of \mathfrak{F}_0 , $\{g_j\}$ for a basis of \mathfrak{F}_1 , and expand $u = u_j f_j$, collecting the coefficients (u_j) into a vector \mathbf{u} . A cardinal basis is one in which $f_j(x_i) = \delta_{ij}$, so that $u_j = u(x_j)$.

Definition 1. A linear operator

$$\mathcal{D} : \mathfrak{F}_0 \times \mathfrak{F}_1 \rightarrow \mathfrak{F}, \quad \mathcal{D}(u)v \mapsto w,$$

is formally skew-adjoint if there is a functional $b(u, v, w)$, depending only on the boundary values of u, v , and w and their derivatives up to a finite order, such that

$$\langle v, \mathcal{D}(u)w \rangle = -\langle w, \mathcal{D}(u)v \rangle + b(u, v, w) \quad \forall u \in \mathfrak{F}_0, \forall v, w \in \mathfrak{F}_1.$$

\mathfrak{F}_1 is called a domain of interior skewness of \mathcal{D} . If $b(u, v, w) = 0 \quad \forall u \in \mathfrak{F}_0, \forall v, w \in \mathfrak{F}_1$, \mathfrak{F}_1 is called a domain of skewness of \mathcal{D} , and we say that \mathcal{D} is skew-adjoint.

Example 1. Let $\mathfrak{F}^{\text{pp}}(n, r) = \{u \in C^r([-1, 1]) : u|_{[x_i, x_{i+1}]} \in \mathfrak{P}_n\}$ be the piecewise polynomials of degree n with r derivatives. For $\mathcal{D} = \partial_x$, $\mathfrak{F}^{\text{pp}}(n, r)$, $n, r \geq 0$, is a domain of interior skewness, i.e., continuity suffices, and $\{u \in \mathfrak{F}^{\text{pp}}(n, r) : u(\pm 1) = 0\}$ is a domain of skewness.

Example 2. With $\mathcal{D}(u) = 2(u\partial_x + \partial_x u) + \partial_{xxx}$, we have

$$\langle v, \mathcal{D}(u)w \rangle + \langle w, \mathcal{D}(u)v \rangle = [w_{xx}v - w_x v_x + w v_{xx} + 2uvw],$$

so suitable domains of interior skewness are $\mathfrak{F}_0 = \mathfrak{F}^{\text{pp}}(1, 0)$, $\mathfrak{F}_1 = \mathfrak{F}^{\text{pp}}(3, 2)$, i.e., more smoothness is required from v and w than from u . A boundary condition which makes $\mathcal{D}(u)$ skew is $\{v : v(\pm 1) = 0, v_x(1) = v_x(-1)\}$.

Definition 2. \mathfrak{F}_0 is natural for \mathcal{H} if $\forall u \in \mathfrak{F}_0$ there exists $\frac{\delta \mathcal{H}}{\delta u} \in \mathfrak{F}$ such that

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{H}(u + \varepsilon v) - \mathcal{H}(u)}{\varepsilon} = \left\langle v, \frac{\delta \mathcal{H}}{\delta u} \right\rangle \quad \forall v \in \mathfrak{F}.$$

The naturality of \mathfrak{F}_0 often follows from the vanishing of the boundary terms, if any, which appear of the first variation of \mathcal{H} , together with mild smoothness assumptions.

We use appropriate spaces \mathfrak{F}_0 and \mathfrak{F}_1 to generate spectral, pseudospectral, and finite element discretizations which have discrete energy $H := \mathcal{H}|_{\mathfrak{F}_0}$ as a conserved quantity. The discretization of the differential operator \mathcal{D} is a linear operator $\overline{\mathcal{D}} : \mathfrak{F}_1 \rightarrow \mathfrak{F}_0$, and the discretization of the variational derivative $\frac{\delta \mathcal{H}}{\delta u}$ is $\frac{\delta \overline{\mathcal{H}}}{\delta u} \in \mathfrak{F}_1$. Each of $\overline{\mathcal{D}}$ and $\frac{\delta \overline{\mathcal{H}}}{\delta u}$ is a weighted residual approximation [3], but each uses spaces of weight functions different from its space of trial functions.

Definition 3. S is the matrix of $\langle \cdot, \cdot \rangle|_{\mathfrak{F}_0 \times \mathfrak{F}_1}$, i.e. $S_{ij} := \langle f_i, g_j \rangle$. $A(u)$ is the matrix of the linear operator $\mathcal{A} : (v, w) \mapsto \langle v, \mathcal{D}(u)w \rangle$, i.e. $A_{ij}(u) := \langle g_i, \mathcal{D}(u)g_j \rangle$.

Proposition 1. Let \mathfrak{F}_0 be natural for \mathcal{H} and let S be nonsingular. Then for every $u \in \mathfrak{F}_0$ there is a unique element $\overline{\frac{\delta \mathcal{H}}{\delta u}} \in \mathfrak{F}_1$ such that

$$\left\langle w, \overline{\frac{\delta \mathcal{H}}{\delta u}} \right\rangle = \left\langle w, \frac{\delta \mathcal{H}}{\delta u} \right\rangle \quad \forall w \in \mathfrak{F}_0.$$

Its coordinate representation is $S^{-1} \nabla H$ where $H(\mathbf{u}) := \mathcal{H}(u_i f_i)$.

Proposition 2. Let S be nonsingular. For every $v \in \mathfrak{F}_1$, there exists a unique element $\overline{\mathcal{D}}v \in \mathfrak{F}_0$ satisfying

$$\langle \overline{\mathcal{D}}v, w \rangle = \langle \mathcal{D}v, w \rangle \quad \forall w \in \mathfrak{F}_1.$$

The map $v \mapsto \overline{\mathcal{D}}v$ is linear, with matrix representation $D := S^{-T}A$.

Definition 4. $\overline{\mathcal{D}} \frac{\delta \mathcal{H}}{\delta u} : \mathfrak{F}_0 \rightarrow \mathfrak{F}_0$ is the dual composition discretization of $\mathcal{D} \frac{\delta \mathcal{H}}{\delta u}$.

Its matrix representation is $S^{-T}AS^{-1}\nabla H$. The name “dual composition” comes from the dual roles played by \mathfrak{F}_0 and \mathfrak{F}_1 in defining $\overline{\mathcal{D}}$ and $\overline{\frac{\delta \mathcal{H}}{\delta u}}$ which is necessary so that their composition has the required linear-gradient structure. Implementation and accuracy of dual composition and Galerkin discretizations are similar. Because they coincide in simple cases, such methods are widely used already.

Proposition 3. If \mathfrak{F}_1 is a domain of skewness, the matrix $S^{-T}AS^{-1}$ is antisymmetric, and the system of ODEs

$$(4) \quad \dot{\mathbf{u}} = S^{-T}AS^{-1}\nabla H$$

has H as an integral. If, in addition, \mathcal{D} is constant—i.e., does not depend on u —then the system (4) is Hamiltonian.

The method of dual compositions also yields discretizations of linear differential operators \mathcal{D} (by taking $\mathcal{H} = \frac{1}{2} \langle u, u \rangle$), and discretizations of variational derivatives (by taking $\mathcal{D} = 1$). It also applies to formally self-adjoint \mathcal{D} ’s and to mixed (e.g. advection-diffusion) operators, where preserving symmetry gives control of the energy.

The composition of two weighted residual discretizations is not necessarily itself of weighted residual type. The simplest case is when $\mathfrak{F}_0 = \mathfrak{F}_1$ and we compare the dual composition to the Galerkin discretization, a weighted residual discretization of $\mathcal{D} \frac{\delta \mathcal{H}}{\delta u}$ with trial functions and weights both in \mathfrak{F}_0 . They are the same when projecting $\frac{\delta \mathcal{H}}{\delta u}$ to \mathfrak{F}_0 , applying \mathcal{D} , and again projecting to \mathfrak{F}_0 , is equivalent to directly projecting $\mathcal{D} \frac{\delta \mathcal{H}}{\delta u}$ to \mathfrak{F}_0 .

For brevity, we assume $\mathfrak{F}_0 = \mathfrak{F}_1$ for the rest of Section 2.

Proposition 4. $\overline{\mathcal{D}} \frac{\delta \mathcal{H}}{\delta u}$ is the Galerkin approximation of $\mathcal{D} \frac{\delta \mathcal{H}}{\delta u}$ if and only if $\mathcal{D} \left(\overline{\frac{\delta \mathcal{H}}{\delta u}} - \frac{\delta \mathcal{H}}{\delta u} \right) \perp \mathfrak{F}_0$. This occurs if (i) $\mathcal{D}(\mathfrak{F}_0^\perp) \perp \mathfrak{F}_0$, or (ii) $\overline{\mathcal{D}}$ is exact and applying \mathcal{D} and orthogonal projection to \mathfrak{F}_0 commute, or (iii) $\overline{\frac{\delta \mathcal{H}}{\delta u}}$ is exact, i.e., $\frac{\delta \mathcal{H}}{\delta u} \in \mathfrak{F}_0$.

Fourier spectral methods with $\mathcal{D} = \partial_x^n$ satisfy (ii), since then \mathfrak{F} has an orthogonal basis of eigenfunctions e^{ijx} of \mathcal{D} , and differentiating and projecting (dropping the high modes) commute. This is illustrated later for the KdV equation.

The most obvious situation in which $\frac{\delta \mathcal{H}}{\delta u} \in \mathfrak{F}_0$ is when $\mathcal{H} = \frac{1}{2} \langle u, u \rangle$, since then $\frac{\delta \mathcal{H}}{\delta u} = u \in \mathfrak{F}_0$ and $\mathcal{D} \frac{\delta \mathcal{H}}{\delta u} = \mathcal{D}u$, and the discretization of \mathcal{D} is obviously the Galerkin one! When the functions f_j are nonlocal, D is often called the spectral differentiation matrix. The link to standard pseudospectral methods is that some Galerkin methods are pseudospectral.

Proposition 5. *If $\mathcal{D}(\mathfrak{F}_1) \subseteq \mathfrak{F}_1$, then $\overline{\mathcal{D}}v = \mathcal{D}v$, i.e., the Galerkin approximation of the derivative is exact. If, further, $\{f_j\}$ is a cardinal basis, then D is the standard pseudospectral differentiation matrix, i.e. $D_{ij} = \mathcal{D}f_j(x_i)$.*

We want to emphasize that although A , S , and D depend on the basis, $\overline{\mathcal{D}}$ depends only on \mathfrak{F}_0 and \mathfrak{F}_1 , i.e., it is basis and grid independent. In the factorization $D = S^{-T}A$, the (anti)symmetry of A and S is basis independent, unlike that of D . These points are well known in finite elements, less so in pseudospectral methods.

Example 3 (Fourier differentiation). Let \mathfrak{F}_1 be the trigonometric polynomials of degree n , which is closed under differentiation (so that Prop. 5) applies, and is a domain of skewness of $\mathcal{D} = \partial_x$. In any basis, A is antisymmetric. Furthermore, the two popular bases, $\{\sin(jx)_{j=1}^n, \cos(jx)_{j=0}^n\}$, and the cardinal basis on equally-spaced grid points, are both orthogonal, so that $S = \alpha I$ and $D = S^{-1}A$ is antisymmetric in both cases.

Example 4 (Polynomial differentiation). $\mathfrak{F}_1 = \mathfrak{P}_n([-1, 1])$ is a domain of interior skewness which is closed under $\mathcal{D} = \partial_x$, so pseudospectral differentiation factors as $D = S^{-1}A$ in any basis. For a cardinal basis which includes $x_0 = -1$, $x_n = 1$, we have $(A + A^T)_{ij} = -1$ for $i = j = 0, 1$ for $i = j = n$, and 0 otherwise, making obvious the influence of the boundary. For the Chebyshev points $x_i = -\cos(i\pi/n)$, $i = 0, \dots, n$, A can be evaluated first in a basis $\{T_i\}$ of Chebyshev polynomials: one finds $A_{ij}^{\text{cheb}} = 2j^2/(j^2 - i^2)$ for $i - j$ odd, and $S_{ij}^{\text{cheb}} = 2(i^2 + j^2 - 1)/[((i + j)^2 - 1)((i - j)^2 - 1)]$ for $i - j$ even, with other entries 0. Changing to a cardinal basis by $F_{ij} = T_j(x_i) = \cos(ij\pi/n)$, a discrete cosine transform, gives $A = F^{-1}A^{\text{cheb}}F^{-T}$. For example, with $n = 3$ (so that $(x_0, x_1, x_2, x_3) = (-1, -\frac{1}{2}, \frac{1}{2}, 1)$), we have

$$D = \frac{1}{6} \begin{pmatrix} -19 & 24 & -8 & 3 \\ 2 & -6 & -2 & 6 \\ -6 & 2 & 6 & -2 \\ -3 & 8 & -24 & 19 \end{pmatrix} = S^{-T}A = \frac{1}{256} \begin{pmatrix} 4096 & -304 & 496 & -1024 \\ -304 & 811 & -259 & 496 \\ 496 & -259 & 811 & -304 \\ -1024 & 496 & -304 & 4096 \end{pmatrix} \frac{1}{270} \begin{pmatrix} -135 & 184 & -72 & 23 \\ -184 & 0 & 256 & -72 \\ 72 & -256 & 0 & 184 \\ -23 & 72 & -184 & 135 \end{pmatrix}.$$

S and A may be more amenable to study than D itself. All their eigenvalues are very well-behaved; none are spurious. The eigenvalues of A are all imaginary and, as $n \rightarrow \infty$, uniformly fill $[-i\pi, i\pi]$ (with a single zero eigenvalue corresponding to the Casimir of ∂_x). The eigenvalues of S closely approximate the quadrature weights of the Chebyshev grid.

For $\mathcal{D} \neq \partial_x$, $\overline{\mathcal{D}}$ may be quite expensive and no longer pseudospectral. (There is in general no S with respect to which the pseudospectral approximation of $\mathcal{D}v$ is skew-adjoint.) However, $\overline{\mathcal{D}}v$ can be computed quickly if fast transforms between cardinal and orthonormal bases exist. We evaluate $\mathcal{D}v$ exactly for $v \in \mathfrak{F}_1$ and then project S -orthogonally to \mathfrak{F}_1 .

Example 5 (Fast Fourier Galerkin method). Let $\mathcal{D}(u)$ be linear in u , for example, $\mathcal{D}(u) = u\partial_x + \partial_x u$. Let $u, v \in \mathfrak{F}_1$, the trigonometric polynomials of degree n . Then

$\mathcal{D}(u)v$ is a trigonometric polynomial of degree $2n$, the first n modes of which can be evaluated exactly using antialiasing and Fourier pseudospectral differentiation. The approximation whose error is orthogonal to \mathfrak{F}_1 is just these first n modes, because $S = I$ in the spectral basis. That is, the antialiased pseudospectral method is here identical to the Galerkin method, and hence skew-adjoint. Antialiasing makes pseudospectral methods conservative. This is the case of the linear \mathcal{D} 's of the Euler fluid equations.

Example 6 (Fast Chebyshev Galerkin method). Let $\mathcal{D}(u)$ be linear in u and let $u, v \in \mathfrak{F}_1 = \mathfrak{P}_n$. With respect to the cardinal basis on the Chebyshev grid with $n + 1$ points, $\overline{\mathcal{D}}(u)v$ can be computed in time $\mathcal{O}(n \log n)$ as follows: (i) Using an FFT, express u and v as Chebyshev polynomial series of degree n ; (ii) Pad with zeros to get Chebyshev polynomial series of formal degree $2n$; (iii) Transform back to a Chebyshev grid with $2n + 1$ points; (iv) Compute the pseudospectral approximation of $\mathcal{D}(u)v$ on the denser grid. Being a polynomial of degree $\leq 2n$, the corresponding Chebyshev polynomial series is exact; (v) Convert $\mathcal{D}(u)v$ to a Legendre polynomial series using a fast transform [1]; (vi) Take the first $n + 1$ terms. This produces $\overline{\mathcal{D}}(u)v$, because the Legendre polynomials are orthogonal. (vii) Convert to a Chebyshev polynomial series with $n + 1$ terms using a fast transform; (viii) Evaluate at the points of the original Chebyshev grid using an FFT.

3. Examples of the dual composition method.

Example 7 (The KdV equation). $\dot{u} + 6uu_x + u_{xxx} = 0$ with periodic boundary conditions has features which can be used to illustrate various properties of the dual composition method. Consider two of its Hamiltonian forms,

$$\dot{u} = \mathcal{D}_1 \frac{\delta \mathcal{H}_1}{\delta u}, \quad \mathcal{D}_1 = \partial_x, \quad \mathcal{H}_1 = \int \left(-u^3 + \frac{1}{2}u_x^2 \right) dx,$$

and

$$\dot{u} = \mathcal{D}_2 \frac{\delta \mathcal{H}_2}{\delta u}, \quad \mathcal{D}_2 = -(2u\partial_x + 2\partial_x u + \partial_{xxx}), \quad \mathcal{H}_2 = \frac{1}{2} \int u^2 dx.$$

In the case $\mathfrak{F}_0 = \mathfrak{F}_1 = \mathfrak{F}^{\text{trig}}$, $v := \overline{\frac{\delta \mathcal{H}_1}{\delta u}}$ is the orthogonal projection to \mathfrak{F}_0 of $\frac{\delta \mathcal{H}_1}{\delta u} = -3u^2 - u_{xx}$; this can be computed by multiplying out the Fourier series and dropping all but the first n modes, or by antialiasing. Then $\overline{\mathcal{D}_1}v = v_x$, since differentiation is exact in $\mathfrak{F}^{\text{trig}}$. Since \mathcal{D}_1 is constant, the discretization is a Hamiltonian system, and since $\overline{\mathcal{D}_1}$ is exact on constants, it also preserves the Casimir $\mathcal{C} = \int u dx$. In this formulation, Prop. 4 (ii) shows that the dual composition and Galerkin approximations of $\mathcal{D}_1 \frac{\delta \mathcal{H}_1}{\delta u}$ coincide, for differentiation does not map high modes to lower modes, i.e., $\mathcal{D}_1(\mathfrak{F}^{\text{trig}\perp}) \perp \mathfrak{F}^{\text{trig}}$.

In the second Hamiltonian form, $H_2 = \frac{1}{2} \mathbf{u}^T S \mathbf{u}$, $\frac{\delta \mathcal{H}_2}{\delta u} = S^{-1} \nabla H_2 = \mathbf{u}$, and the Galerkin approximation of $\frac{\delta \mathcal{H}_2}{\delta u}$ is exact, so that Prop. 4 (iii) implies that the composition $\overline{\mathcal{D}_2} \frac{\delta \mathcal{H}_2}{\delta u}$ also coincides with the Galerkin approximation. $\overline{\mathcal{D}_2}v$ can be evaluated using antialiasing as in Example 5. $\overline{\mathcal{D}_2}$ is not a Hamiltonian operator, but still generates a skew-gradient system with integral H_2 . Thus in this (unusual) case, the Galerkin and antialiased pseudospectral methods coincide and have three conserved quantities, H_1 , H_2 , and $\mathcal{C}|_{\mathfrak{F}^{\text{trig}}}$.

The situation for finite element methods with $\mathfrak{F}_0 = \mathfrak{F}_1 = \mathfrak{F}^{\text{pp}}(n, r)$ is different. In the first form, we need $r \geq 1$ to ensure that \mathfrak{F}_0 is natural for \mathcal{H}_1 ; in the second form, naturality is no restriction, but we need $r \geq 2$ to ensure that \mathfrak{F}_1 is a domain of interior

skewness. The first dual composition method is still Hamiltonian with integral H_1 and Casimir $C = u_i \int f_i dx$, but because $\overline{\mathcal{D}}_1$ does not commute with projection to \mathfrak{F}_1 , it is *not* a standard Galerkin method. In the second form, $\frac{\delta \mathcal{H}_2}{\delta u} = u$ is still exact, so the dual composition and Galerkin methods still coincide. However, they are not Hamiltonian.

Example 8 (An inhomogeneous wave equation). When natural and skew boundary conditions conflict, it is necessary to take $\mathfrak{F}_0 \neq \mathfrak{F}_1$. Consider $\dot{q} = a(x)p$, $\dot{p} = q_{xx}$, $q_x(\pm 1, t) = 0$. This is a canonical Hamiltonian system with

$$\mathcal{D} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathcal{H} = \frac{1}{2} \int_{-1}^1 (a(x)p^2 + q_x^2) dx, \quad \frac{\delta \mathcal{H}}{\delta q} = -q_{xx}, \quad \frac{\delta \mathcal{H}}{\delta p} = a(x)p.$$

Note that (i) the boundary condition is natural for \mathcal{H} , and (ii) no boundary conditions are required for \mathcal{D} to be skew-adjoint in L^2 . Since $\frac{\delta \mathcal{H}}{\delta u}$ is computed with trial functions in \mathfrak{F}_1 , we should not include $q_x(\pm 1) = 0$ in \mathfrak{F}_1 , for this would be to enforce $(-q_{xx})_x = 0$. In [8] we show that a spectrally accurate dual composition method is obtained with $\mathfrak{F}_0 = \{q \in \mathfrak{P}_{n+2} : q_x(\pm 1) = 0\} \times \mathfrak{P}_n$ and $\mathfrak{F}_1 = \mathfrak{P}_n \times \mathfrak{P}_n$.

4. Quadrature of Hamiltonians. Computing $\nabla H = \nabla \mathcal{H}(u_j f_j)$ is not always possible in closed form. We would like to approximate \mathcal{H} itself by quadratures in real space. However, even if the discrete H and its gradient are spectrally accurate approximations, they cannot always be used to construct spectrally accurate Hamiltonian discretizations.

In a cardinal basis, let $\mathcal{H} = \int h(u) dx$ and define the quadrature Hamiltonian $H_q := h(u_j) w_j = \mathbf{w}^T h(\mathbf{u})$ where $w_j = \int f_j dx$ are the quadrature weights. Since $\nabla H_q = W h'(\mathbf{u})$, $\frac{\delta \mathcal{H}}{\delta u} \approx W^{-1} \nabla H_q$. Unfortunately, $DW^{-1} \nabla H_q$ is not a skew-gradient system, while $DS^{-1} \nabla H_q$ is skew-gradient, but is not an accurate approximation.

$DW^{-1} \nabla H_q$ can only be a skew-gradient system if DW^{-1} is antisymmetric, which occurs in three general cases. (i) On a constant grid, W is a multiple of the identity, so if D is antisymmetric, DW^{-1} is too. (ii) On an arbitrary grid with $D = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, DW^{-1} is antisymmetric. (iii) On a Legendre grid with $\mathfrak{F}_0 = \mathfrak{F}_1$, $S = W$, and $DW^{-1} = W^{-1}AW^{-1}$ is antisymmetric. The required compatibility between D and W remains an intriguing and frustrating obstacle to the systematic construction of conservative discretizations of strongly nonlinear PDEs.

REFERENCES

- [1] B. Alpert and V. Rokhlin. A fast algorithm for evaluation of Legendre expansions. *SIAM J. Sci. Comput.*, 12:158–179, 1991.
- [2] M. H. Carpenter and D. Gottlieb. Spectral methods on arbitrary grids. *J. Comput. Phys.*, 129:74–86, 1996.
- [3] B. A. Finlayson. *Method of Weighted Residuals and Variational Principles*. Academic Press, New York, 1972.
- [4] H.-O. Kreiss and G. Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. In C. de Boor, editor, *Mathematical Aspects of Finite Elements in Partial Differential Equations: Symposium 1974*, pages 195–212. Academic Press, 1974.
- [5] R. I. McLachlan. Spatial discretization of partial differential equations with integrals, preprint.
- [6] R. I. McLachlan. Symplectic integration of Hamiltonian wave equations. *Numer. Math.*, 66:465–492, 1994.

- [7] R. I. McLachlan, G.R.W. Quispel, and N. Robidoux. Unified approach to Hamiltonian systems, Poisson systems, gradient systems, and systems with Lyapunov functions or first integrals. *Phys. Rev. Lett.*, 81(12):2399–2402, 1998.
- [8] R. I. McLachlan and N. Robidoux. Antisymmetry, pseudospectral methods, weighted residual discretizations, and conservative PDEs. *J. Comput. Phys.*, submitted.
- [9] Pelle Olsson. Summation by parts, projections, and stability. I. *Math. Comput.*, 64:1035–1065, S23–S26, 1995.